Appl. Math. Inf. Sci. **8**, No. 2, 681-689 (2014)

681

Recognition of Concurrent Control Chart Patterns by Integrating ICA and SVM

Chi-Jie Lu[1], Yuehjen E. Shao[2,] and Chin-Chan Li[2]*

[1] Department of Industrial Management, Chien Hsin University of Science and Technology, Taiwan, ROC
[2] Department of Statistics and Information Science, Fu Jen Catholic University, Taiwan, ROC

Abstract: Control chart patterns (CCPs) recognition is an important issue in statistical process control, since abnormal CCPs exhibited in control charts can be associated with certain assignable causes and then help to quickly eliminate assignable causes of process variation. Most of the existing studies assume that the observed process data needs to be recognized are basic types of abnormal CCPs. However, in practical situations, the observed process data may be concurrent CCPs which are mixed together by two basic CCPs. In this study, an integrated scheme using independent component analysis (ICA) and support vector machine (SVM) is proposed for recognizing concurrent CCPs. The proposed ICA-SVM scheme initially uses ICA with concurrent patterns for generating independent components (ICs). The hidden basic patterns of the concurrent patterns can be discovered in these ICs. The ICs are then used as input variables for the SVM for building a CCP recognition model. Experimental results reveal that the proposed ICA-SVM scheme can produce accurate and stable recognition results. It outperforms the three comparison models and is able to effectively recognize concurrent control chart patterns.

Keywords: Control chart pattern recognition, concurrent control chart pattern, independent component analysis, support vector machine.

1 Introduction

Control charts are important statistical process control tools for determining whether a process is in control or out of control. A process is in control when only common cause variability exists. A process is out of control when a data point falls outside the control limits or a series of data points exhibit unnatural/abnormal patterns [1]. If a process is out of control, corrective action should be taken to search for and eliminate assignable causes of variation. Since unnatural control chart patterns (CCPs) associate a specific set of assignable causes in manufacturing processes, effective recognition of unnatural CCPs is an important issue in SPC. Five basic CCPs are commonly exhibited in control charts including normal (NOR), systematic (SYS), cyclic (CYC), trend (TRE) and shift (SHI) [2,3]. Except for the normal pattern, all other patterns are abnormal CCPs and indicate that the process

being monitored is out of control and requires adjustment. Figure 1 shows these five basic control chart patterns.

There have been many approaches proposed for CCP recognition [3,4,5]. Most existing studies are concerned with the recognition of single abnormal control chart patterns (as shown in Figure 1). However, in real applications, concurrent patterns, where two single patterns exist together, are frequently encountered, which may be associated with different assignable causes. When a concurrent pattern occurs, most existing approaches will classify the input pattern into one of predefined single CCPs. Unfortunately, single CCPs cannot represent the characteristics of the concurrent patterns and further tracking of its assignable causes also becomes impossible. Concurrent patterns are more difficult to recognize than single patterns due to the complexity of the pattern interaction. Without loss of generality, Figure 2 shows six concurrent CCPs which are mixed by two basic patterns. Compared to the patterns illustrated in Figure 1, it can be

* Corresponding author e-mail: stat1003@mail.fju.edu.tw
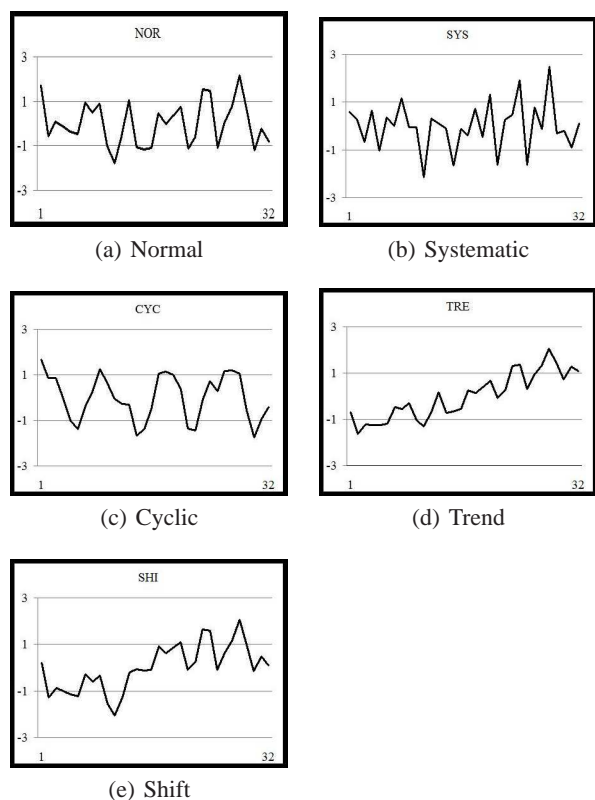
Fig. 1. Five basic control chart patterns.
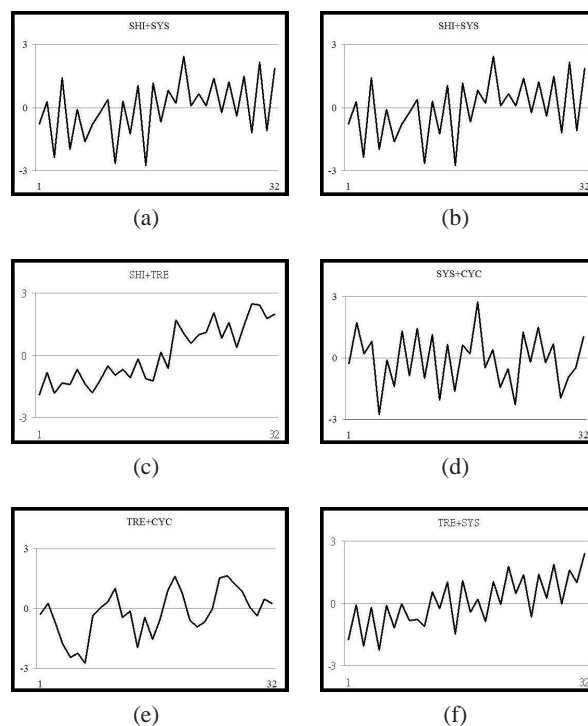
Fig. 2. Six concurrent CCPs: (a) Shift+Cyclic; (b) Shift+Systematic; (c) Shift+Trend; (d) Systematic+Cyclic; (e) Trend+Cyclic; (f) Trend+Systematic.

observed from Figure 2 that the concurrent CCPs are more difficult to recognize than the basic CCPs.

Only few studies have addressed the recognition of concurrent CCPs. Guh and Tannock [6] used the back-propagation neural network (BPN) to recognize concurrent CCPs. Chen et al. [7] integrated wavelet method and back-propagation neural network for on-line recognition of mixture CCPs. Wang et al. [8] proposed an integrated approach by combining wavelet analysis and neural network for the recognition of concurrent CCPs. These results indicated that neural network-based approaches for recognizing concurrent CCPs can perform well.

Departing from previous approaches, our novel hybrid approach for concurrent CCPs recognition is explored by integrating independent component analysis (ICA) and support vector machine (SVM). ICA is a novel feature extraction technique to find independent sources given only observed data that are mixtures of the unknown sources, without any prior knowledge of the mixing mechanisms [9]. The independent sources, called independent components (ICs), are hidden information of the observable data. ICA has been employed successfully in various fields of multivariate data processing, from signal processing to time series prediction [9]. However,

there are still few applications of using ICA in control chart pattern recognition. Lu et al. [10] integrated ICA, engineering process control and BPN to recognize shift and trend patterns in correlated process. Lu [11] proposed an ICA-based monitoring scheme to identify shift pattern in an autocorrelated process. Lu et al. [12] combined ICA and SVM for diagnosing mixture CCPs which are mixed by the normal and other abnormal basic patterns. However, their work did not consider the issue of recognition of concurrent CCPs.

Support vector machine (SVM), based on statistical learning theory, is a novel neural network algorithm [13]. It can lead to great potential and superior performance in practical applications. This is largely due to the structure risk minimization principles in SVM, which has greater generalization ability and is superior to the empirical risk minimization principle as adopted in neural networks. The SVM has attracted the interest of researchers and has been applied many applications such as diseases classification and process monitoring [14,15,16,17,18]. However, relatively few studies have been conducted using SVM for CCP recognition [12].

The proposed ICA-SVM scheme first uses ICA for the observed process data contained mixture patterns for generating independent components. The estimated ICs

are then served as the independent sources of the mixture patterns. The hidden basic patterns of the mixture patterns could be discovered in these ICs. The ICs are then used as the input variables of the SVM for building CCP recognition model.

The rest of this paper is organized as follows. Section 2 gives brief overviews of ICA and SVM. The proposed model is described in Section 3. Section 4 presents the experimental results and this study is concluded in Section 5.

2 Methodology

2.1 Independent component analysis

In the basic conceptual framework of ICA algorithm [9], it is assumed that m measured variables, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]^T$ be a multivariate data matrix of size $m \times n$, $m \le n$, consisting of observed mixture signals $\mathbf{x}_i$ of size $1 \times n$, $i = 1, 2, \ldots, m$. In the basic ICA model, the matrix $\mathbf{X}$ can be modeled as [9]

$$\mathbf{X} = \mathbf{AS} = \sum_{i=1}^{m} \mathbf{a}_i \mathbf{s}_i, \qquad (1)$$

where $\mathbf{a}_i$ is the ith column of the $m \times m$ unknown mixing matrix $\mathbf{A}$; $\mathbf{s}_i$ is the ith row of the $m \times n$ source matrix $\mathbf{S}$. The vectors $\mathbf{s}_i$ are latent source signals that cannot be directly observed from the observed mixture signals $\mathbf{x}_i$. The ICA model aims at finding an $m \times m$ de-mixing matrix $\mathbf{W}$ such that

$$\mathbf{Y} = [\mathbf{y}_i] = \mathbf{WX}, \qquad (2)$$

where $\mathbf{y}_i$ is the ith row of the matrix $\mathbf{Y}$, $i = 1, 2, \ldots, m$. For using vectors $\mathbf{y}_i$ to estimate the independent latent source signals ($\mathbf{s}_i$), $\mathbf{y}_i$ must be statistically independent, and are called independent components (ICs). When de-mixing matrix $\mathbf{W}$ is the inverse of mixing matrix $\mathbf{A}$, i.e. $\mathbf{W} = \mathbf{A}^{-1}$, ICs ($\mathbf{y}_i$) can be used to estimate the latent source signals $\mathbf{s}_i$.

The ICA modeling is formulated as an optimization problem by setting up the measure of the independence of ICs as an objective function and using some optimization techniques for solving the de-mixing matrix $\mathbf{W}$. Several existing algorithms can be used to perform ICA modeling [9]. In general, the ICs are obtained by using the de-mixing matrix $\mathbf{W}$ to multiply the matrix $\mathbf{X}$, i.e. $\mathbf{Y} = \mathbf{WX}$. The de-mixing matrix $\mathbf{W}$ can be determined using an unsupervised learning algorithm with the objective of maximizing the statistical independence of ICs. And the statistical independence of ICs can be measured in terms of their non-Gaussian properties.

Normally, non-Gaussianity can be verified by two common statistics: kurtosis and negentropy. The kurtosis

of a random variable $\mathbf{y}$, fourth-order cumulant, is classically defined by

$$kurt(\mathbf{y}) = E(\mathbf{y}^4) - 3(E(\mathbf{y}^2))^2 \qquad (3)$$

If variable $\mathbf{y}$ is assumed to be zero mean and unit variance, the right-hand side simplifies to $E(\mathbf{y}^4) - 3$. This shows that kurtosis is simply a normalized version of the fourth moment $E(\mathbf{y}^4)$. For a Gaussian $\mathbf{y}$, the fourth moment equals $3(E(\mathbf{y}^2))^2$. Thus, kurtosis is zero for a Gaussian random variable and non-zero for most non-Gaussian random variables.

Unlike kurtosis, negentropy is determined according to the information quantity of (differential) entropy. Entropy is a measure of the average uncertainty in a random variable. The differential entropy H of random variable $\mathbf{y}$ with density $p(\mathbf{y})$ defined as $H(\mathbf{y}) = -\int p(\mathbf{y}) \log p(\mathbf{y}) d\mathbf{y}$. According to a fundamental result of information theory, a Gaussian variable will have the highest entropy value among a set of random variables with equal variance [9]. For obtaining a measure of non-Gaussianity, the negentropy J is defined as follows:

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}) \qquad (4)$$

where $\mathbf{y}_{gauss}$ is a Gaussian random vector having the same covariance matrix as $\mathbf{y}$.

The negentropy is always non-negative and is zero if and only if $\mathbf{y}$ has a Gaussian distribution. Since the problem in using negentropy is computationally very difficult, an approximation of negentropy is proposed as follows:

$$J(y) \approx [E\{G(y)\} - E\{G(v)\}]^2 \qquad (5)$$

where v is a Gaussian variable of zero mean and unit variance, and y is a random variable with zero mean and unit variance. G is a nonquadratic function, and is given by $G(y) = \exp(-y^2/2)$ in this study. The *FastICA* algorithm proposed by [9] is adopted in this paper to solve for the de-mixing matrix $\mathbf{W}$.

The ICA model in Eq.(1) shows the limitation of the ICA model that is, one cannot determine the order of the ICs because of the simultaneous unknowns of $\mathbf{S}$ and $\mathbf{A}$. Hyvärinen et al. [9] suggested that ICs can be sorted according to their non-Gaussianity. In this paper, the ICs are sorted based on their kurtosis values.

2.2 Support vector machine

The basic idea of applying SVM to pattern recognition can be stated briefly as follows. We can initially map the input vectors into one feature space (possible with a higher dimension), either linearly or non-linearly, which is relevant with the selection of the kernel function. Then, within the feature space from the first, we seek an

optimized linear division, that is, construct a hyperplane which separates two classes (this can be extended to multi-class).

A description of SVM algorithm is follows. Let $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, $\mathbf{x}_i \in \mathbb{R}^d$, $y_i \in \{-1, 1\}$ be the training set with input vectors and labels. Here, N is the number of sample observations and d is the dimension of each observation, y_i is the known target. The algorithm is to seek for the hyperplane $\mathbf{w} \cdot \mathbf{x}_i + b = 0$, where $\mathbf{w}$ is the vector of hyperplane and b is a bias term, to separate the data from two classes with maximal margin width $2/\|\mathbf{w}\|^2$, and the all the points under the boundary is named support vector. The hyperplane can be optimized using SVM by solving the following optimization problem [13].

$$\text{Min } \Phi(\mathbf{x}) = \frac{1}{2}\|\mathbf{w}\|^2$$
$$\text{S.t. } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \ldots, N. \quad (6)$$

Since it is difficult to solve Eq. (6) directly, then Eq. (6) must be transformed into a dual problem by Lagrange method. The value of in the Lagrange method must be non-negative real coefficients. Problem (6) is transformed into the following dual problem.

$$\text{Max } \Phi(\mathbf{w}, b, \xi, \alpha, \beta) = \sum_{i=1}^N \alpha_i - \frac{1}{2}\sum_{\substack{i=1 \\ j=1}}^N \alpha_i \alpha_j y_i y_j \mathbf{x}_i \mathbf{x}_j$$
$$\text{S.t. } \sum_{j=1}^N \alpha_j y_j = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, N. \quad (7)$$

In Eq. (7), C is the penalty factor and determines the degree of penalty assigned to an error. It can be viewed as a tuning parameter which can be used to control the trade-off between maximizing the margin and the classification error.

In general, it could not find the linear separate hyperplane in all application data. In the non-linear data, transforming the original data to higher dimension of linear separate is the best solution. The higher dimension is called feature space which improves the data separated by classification. The common kernel function are linear, polynomial, radial basis function (RBF) and sigmoid. Although several choices for the kernel function are available, the most widely used kernel function is the RBF kernel defined as $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, $\gamma \geq 0$ [13], where γ denotes the width of the RBF. Thus, the RBF is applied in this study. The original SVM was designed for binary classifications. Constructing multi-class SVM is still an ongoing research issue. In this study, we used multi-class SVM method proposed by [14]. For more details, please refer to [14].

3 The Proposed Integrating Recognition Scheme

This study combines ICA and SVM for recognizing concurrent control chart patterns. Figure 3 shows the research scheme of the proposed ICA-SVM model. As shown in Figure 3, the proposed scheme consists of two stages. In the training stage, the aim is to find the best parameter setting to train SVM model for concurrent CCP recognition. The first step of the training stage is to generate six concurrent CCPs as shown in Figure 2. Then, they are used as training samples to build SVM model. Since the RBF kernel function is adopted in this study, the performance of SVM is mainly affected by the setting of parameters of two parameters i.e. C and γ [19]. There are no general rules for the choice of the parameters. In this study, the grid search proposed by [19] is used in this study for parameters setting. The trained SVM model with best parameter set is preserved and used in the monitoring stage for concurrent CCP recognition.
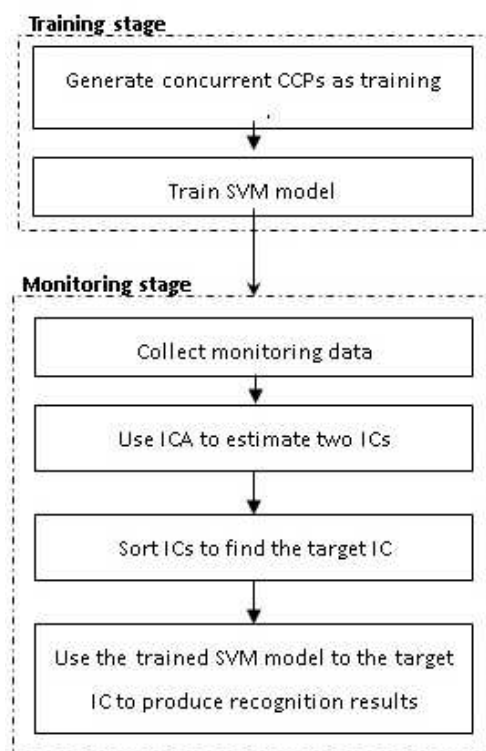
**Fig. 3.** The research scheme of the proposed ICA-SVM model.

In the monitoring stage, the first step is to collect observed data from monitoring multivariate process. Then the collected monitoring data is combined with a

generated pseudo variable which contains normal pattern to produce an input data matrix for performing ICA. This design is based on the assumption of ICA that the dimension of input variables is needed more than two. After using ICA to the input data matrix, two ICs are estimated. In order to find the target IC containing unnatural pattern, the two ICs are sorted based on their non-Gaussianality by kurtosis values. As the kurtosis value of a normal distributed variable is equal to zero, the IC contains the largest kurtosis value is considered as the target IC. Finally, the trained SVM is utilized to the target IC to produce CCP recognition results.

As an example, Figure 4(a) shows the observed data collected from the monitoring process and Figure 4(b) depicts the pseudo variable containing normal pattern. It is assumed that observed data are mixed by the shift and systematic patterns. Then, the ICA model is used to the data to generate two ICs which are illustrated in Figures 4(c) and 4(d). It is clear that Figure 4(c) can be used represent concurrent pattern containing systematic and cyclic patterns and considered as the target IC. Figure 4(d) represents the normal pattern. The trained SVM model is used to the target IC to recognize the pattern exhibited in the IC. According to the SVM results, the process monitoring task is conducted to identify which basic patterns are exhibited in the process.

4 Experimental Results

In this study, five basic CCPs (as shown in Figure 1) are used to construct six concurrent CCPs which are used for
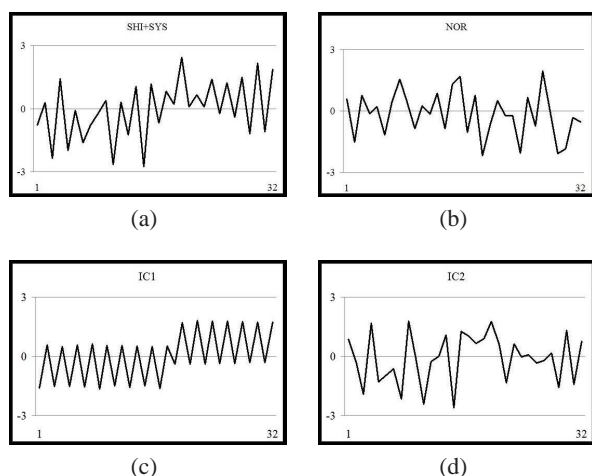
Fig. 4. (a) the observed concurrent data; (b) the pseudo variable containing normal pattern; (c) the IC represents the concurrent pattern containing the shift and systematic patterns; (d) the IC represents the normal pattern.

training and testing the proposed ICA-SVM scheme. The five basic patterns are generated using the same equations and values of different pattern parameters, as used by [3]. The parameters along with the equations used for simulating the basic CCPs are given in Table 1. The values of different parameters for abnormal patterns are randomly varied in a uniform manner between the limits. It is assumed that, in the current approach for pattern generation, all the patterns in an observation window are complete. The observation window used in this study is 32 data points.

Based on the equations of basic CCPs, the equations of generating concurrent CCPs are shown in Table 2. The parameters setting of the concurrent CCPs are the same as that in Table 1.

Features of observation window, such as mean, standard deviation, kurtosis and slop characteristics are adopted in literatures to improve the performance of CCP recognition [3]. However, features may ignore some useful information contained in the original data points of observation window since they are abstract information of the data points. Therefore, the proposed ICA-SVM model directly uses the 32 data pints of observation window as inputs of the SVM model. That is, there are 32 input variables used in the proposed model for CCP recognition. In order to demonstrate the performance of the proposed ICA-SVM scheme, the single SVM model without using ICA as preprocessing is constructed. It also directly uses the data pints of observation window as inputs.

After using the grid search method to the two models, the best parameter sets for the ICA-SVM and single SVM models are ($C = 2^{-1}, \gamma = 2^{-15}$) and ($C = 2^{-2}, \gamma = 2^{-15}$), respectively. Note that the model selection details of the three models are omitted for saving space. The recognition results of the ICA-SVM and single SVM models are respectively illustrated in Table 3 and Table 4.

From Table 3 and Table 4, it can be found that the average correct classification rates of the proposed ICA-SVM model and the single SVM model are 97.23% and 62.98%, respectively. The proposed model outperforms the single SVM model in all testing mixture CCPs. Therefore, the proposed ICA-SVM scheme can effectively recognize concurrent control chart patterns mixed by two basic patterns.

To further evaluate the recognition performance of the proposed ICA-SVM model, a novel neural network algorithm called extreme learning machine (ELM) is also used in this study. The design of the ELM proposed by Huang et al. [20] is based on a single-hidden layer feedforward neural network (SLFNN) with a wide variety of hidden nodes, which randomly generates hidden parameters before analytically determining the output weights [21,22]. Prior investigations have reported that the ELM tends to provide better generalization

Table 1. Parameters for simulating basic control chart patterns

CCPs	Pattern equations	Pattern parameters
NOR	$x(t) = u + \gamma(t)\sigma$	Mean$(u) = 0$ Standard deviation$(\sigma) = 1$
SYS	$x(t) = u + \gamma(t)\sigma + d(-1)^t$	Systematic departure$(d) = (1(\sigma) \text{ to } 3(\sigma))$
CYC	$x(t) = u + \gamma(t)\sigma + a \cdot \sin(2\pi t/p)$	Amplitude$(a) = (1.5(\sigma) \text{ to } 2.5(\sigma))$ Period$(p) = (8 \text{ and } 16)$
TRE	$x(t) = u + \gamma(t)\sigma + ig$	Gradient$(g) = (0.05(\sigma) \text{ to } 0.1(\sigma))$
SHI	$x(t) = u + \gamma(t)\sigma + ks$ $k = 1$ if $t > L$, else $k = 0$.	Shift magnitude$(s) = (1.5(\sigma) \text{ to } 2.5(\sigma))$ Shift position$(L) = (7, 13, 19)$

Table 2. Equations for simulating concurrent control chart patterns

Control chart patterns	Pattern equations
Shift+Cyclic	$x(t) = u + \gamma(t)\sigma + ks + a \cdot \sin(2\pi t/p)$
Shift+Systematic	$x(t) = u + \gamma(t)\sigma + ks + d(-1)^t$
Shift+Trend	$x(t) = u + \gamma(t)\sigma + ks + ig$
Systematic+Cyclic	$x(t) = u + \gamma(t)\sigma + d(-1)^t + a \cdot \sin(2\pi t/p)$
Trend+Cyclic	$x(t) = u + \gamma(t)\sigma + ig + a \cdot \sin(2\pi t/p)$
Trend+Systematic	$x(t) = u + \gamma(t)\sigma + ig + d(-1)^t$

Table 3. Confusion matrix of testing results using the proposed ICA-SVM model.

True pattern class	Identified patterns class					
	Shift+Trend	Shif+Cyc	Shift+Sys	Trend+Cyc	Trend+Sys	Sys+Cyc
Shift+Trend	92.97%	0.00%	7.03%	0.00%	0.00%	0.00%
Shift+Cyc	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Shift+Sys	0.00%	0.00%	98.44%	0.00%	1.56%	0.00%
Trend+Cyc	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%
Trend+Sys	0.00%	0.00%	2.34%	0.00%	97.66%	0.00%
Sys+Cyc	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Average	97.23%					

performance and a much faster learning speed than SVM model [20]. Therefore, the single ELM model using the original data as direct inputs and the ICA-ELM model utilizing the same procedure as the ICA-SVM model described earlier are used as comparison models in this study.

It is known that the most important and critical parameter of ELM is the number of hidden nodes and ELM tends to be unstable in a single run forecasting [21, 22]. Therefore, the single ELM and ICA-ELM models with different numbers of hidden nodes varying from 1 to 30 were constructed. For each number of nodes, the single ELM and ICA-ELM models are repeated 30 times and the number of hidden nodes that gives the smallest testing RMSE value is selected for the single ELM and ICA-ELM models. In this study, the ELM model with 30 hidden nodes and the ICA-ELM model using 21 hidden nodes have smaller RMSE values. Therefore, they are the best models of the ELM and ICA-ELM models, respectively

Table 5 and Table 6 show respectively the recognition results of the ICA-ELM and single ELM models. The average correct classification rates of the ICA-ELM and single ELM models are respectively 54.44% and 21.63%. Compared to the recognition results of the ICA-SVM model, we find that the ICA-ELM and ELM models cannot perform well for recognizing concurrent CCPs. Thus, the ICA-SVM model still has the highest average correct classification rate.

In order to assess the robustness of the ICA-SVM method, the performance of the single SVM, single ELM, ICA-ELM and ICA-SVM models was tested using 10 independent runs. Based on the findings in Table 7, it can be observed that the ICA-SVM method outperforms the three comparison models under all runs. This indicates that the ICA-SVM approach indeed provides better classification accuracy than the other three approaches. Nevertheless, the ICA-SVM method has the lowest standard deviation (S.D.) and hence can provide a more stable classification result. It is worth to note that the recognition performances of the ICA-ELM and

Table 4. Confusion matrix of testing results using the single SVM model.

True pattern class	Identified patterns class					
	Shift+Trend	Shif+Cyc	Shift+Sys	Trend+Cyc	Trend+Sys	Sys+Cyc
Shift+Trend	37.11%	16.41%	35.94%	4.69%	1.56%	0.00%
Shift+Cyc	0.00%	96.88%	0.00%	3.13%	0.00%	0.00%
Shift+Sys	0.00%	0.00%	75.00%	0.78%	24.22%	0.00%
Trend+Cyc	0.00%	50.00%	0.00%	50.00%	0.00%	0.00%
Trend+Sys	0.00%	3.13%	7.81%	11.72%	77.34%	0.00%
Sys+Cyc	0.00%	17.19%	3.13%	1.56%	6.25%	71.88%
Average	**62.98%**					

Table 5. Confusion matrix of testing results using the ICA-ELM model.

True pattern class	Identified patterns class					
	Shift+Trend	Shif+Cyc	Shift+Sys	Trend+Cyc	Trend+Sys	Sys+Cyc
Shift+Trend	76.56%	14.06%	7.42%	0.00%	0.39%	1.56%
Shift+Cyc	8.59%	33.59%	14.84%	35.94%	5.47%	1.56%
Shift+Sys	16.41%	23.44%	30.47%	16.41%	13.28%	0.00%
Trend+Cyc	29.69%	14.84%	0.00%	46.88%	8.59%	0.00%
Trend+Sys	14.06%	2.34%	7.81%	0.00%	75.78%	0.00%
Sys+Cyc	15.63%	17.19%	9.38%	15.63%	14.06%	28.13%
Average	**54.45%**					

Table 6. Confusion matrix of testing results using the single ELM model.

True pattern class	Identified patterns class					
	Shift+Trend	Shif+Cyc	Shift+Sys	Trend+Cyc	Trend+Sys	Sys+Cyc
Shift+Trend	37.89%	23.44%	12.89%	4.69%	17.58%	1.17%
Shift+Cyc	33.59%	21.09%	19.53%	7.81%	16.41%	1.56%
Shift+Sys	38.28%	21.88%	12.50%	10.16%	14.06%	3.13%
Trend+Cyc	37.50%	17.19%	21.09%	7.03%	15.63%	1.56%
Trend+Sys	28.91%	17.97%	21.09%	8.59%	21.88%	1.56%
Sys+Cyc	28.13%	23.44%	20.31%	7.81%	15.63%	4.69%
Average	**21.63%**					

ICA-SVM models are, respectively, better that that of the single ELM and SVM models. This indicates that the ICA model can effectively extract useful information from the observed concurrent CCPs and can be used to further improve the recognition accuracies of the SVM and ELM models.

5 Conclusions

Effective recognition of concurrent CCPs is an important and challenging task in statistical process control. In this study, a scheme for recognizing concurrent CCPs by integrating ICA and SVM is proposed. The proposed scheme, firstly, uses ICA to the observed concurrent pattern to generate ICs. Then, the SVM model is used for the target IC for pattern recognition. Six concurrent CCPs are used in this study for evaluating the performance of the proposed method.

Experimental results show that the proposed ICA-SVM scheme can produce accurate and stable recognition results and outperforms the single SVM, single ELM and ICA-ELM models. According to the results, it can be concluded that the ICA-SVM scheme is an effective alternative to recognize concurrent control chart patterns in manufacturing processes.

Acknowledgements

Table 7. Robustness evaluation of the ICA-SVM, single SVM, ICA-ELM and single ELM models.

Runs / Methods	ICA-SVM	Single SVM	ICA-ELM	Single ELM
1	97.23%	62.98%	54.45%	21.63%
2	95.56%	69.83%	59.13%	29.20%
3	97.00%	59.13%	60.94%	23.68%
4	96.15%	59.74%	53.49%	24.16%
5	94.11%	60.22%	49.40%	31.43%
6	93.75%	60.34%	51.44%	23.08%
7	94.59%	63.94%	52.52%	24.40%
8	94.59%	61.54%	43.87%	24.88%
9	99.76%	61.66%	37.86%	29.63%
10	92.07%	63.70%	52.52%	32.60%
Average	**95.48%**	**62.31%**	**51.56%**	**26.47%**
S.D.	**0.022**	**0.031**	**0.068**	**0.039**

References

[1] D. C. Montgomery, Introduction to Statistical Quality Control. John Wiley & Sons, New York, (2001).

[2] Western Electric., Statistical Quality Control Handbook. Western Electric Company, Indianapolis, (1958).

[3] S. K. Gauri and S. Charkaborty, Recognition of control chart patterns using improved selection of features, Computer & Industrial Engineering, **56**, 1577-1588 (2009).

[4] K. Assaleh and Y. Al-assaf, Feature extraction and analysis for classifying causable patterns in control charts, Computer & Industrial Engineering, **49**, 168-181 (2005).

[5] R. S. Guh, A hybrid learning-based model for on-line detection and analysis of control chart patterns, Computer & Industrial Engineering, **49**, 35-62 (2005).

[6] R. S. Guh and J. D. T. Tannock, Recognition of control chart concurrent patterns using a neural network approach, International Journal of Production Research, **37**, 1743-1765 (1999).

[7] Z. Chen, S. Lu and S. Lam, A hybrid system for SPC concurrent pattern recognition, Advanced Engineering Informatics, **21**, 303-310 (2007).

[8] C. H. Wang, W. Kuo and H. Qi, An integrated approach for process monitoring using wavelet analysis and competitive neural network, International Journal of Production Research, **45**, 227-244 (2007).

[9] A. Hyvärinen, J. Karhunen and E. Oja, Independent Component Analysis. John Wiley & Sons, New York, (2001).

[10] C. J. Lu, C. M. Wu, C. J. Keng and C. C. Chiu, Integrated application of SPC/EPC/ICA and neural networks, International Journal of Production Research, **46**, 873-893 (2008).

[11] C. J. Lu, An independent component analysis-based disturbance separation scheme for statistical process monitoring, Journal of Intelligent Manufacturing, DOI: 10.1007/s10845-010-0394-3, (2010).

[12] C. J. Lu, Y. E. Shao and B. S. Li., Mixture control chart patterns recognition using independent component analysis and support vector machine, Neurocomputing, **74**, 1908-1914 (2011).

[13] V. N. Vapnik, The Nature of Statistical Learning Theory. Springer, Berlin, (2000).

[14] C. W. Hsu and C. J. Lin, A comparison of methods for multiclass support vector machines, IEEE Transactions on Neural Network, **13**, 415-425 (2002).

[15] Y. E. Shao, C. J. Lu and C. C. Chiu, A fault detection system for an autocorrelated process using SPC/EPC/ANN and SPC/EPC/SVM scheme, International Journal of Innovative Computing, Information and Control, **7**, 5417-5428 (2011).

[16] Y. E. Shao and B. S. Hsu, Determining the contributors for a multivariate SPC chart signal using artificial neural networks and support vector machine, International Journal of Innovative Computing, Information and Control, **5**, 4899-4906 (2009).

[17] Y. E. Shao and C. D. Hou, Change point determination for a multivariate process using a two-stage hybrid scheme, Applied Soft Computing, **13**, 1520-1527 (2013).

[18] C. D. Hou and Y. E. Shao, Integrated use of statistical-based approaches and machine learning techniques for tumors classification uing microarray data, Journal of Biobased Materials and Bioenergy, to appear.

[19] C. W. Hsu, C. C. Chang and C. J. Lin, A practical guide to support vector classification, Department of Computer Science and Information Engineering, National Taiwan University Taipei, Taiwan, (2003).

[20] G. B. Huang, Q. Y. Zhu and C. K. Siew, Extreme learning machine: theory and applications, Neurocomputing, **70**, 489-501 (2006).

[21] Y. Lan, Y. C. Soh and G. B. Huang, Constructive hidden nodes selection of extreme learning machine for regression, Neurocomputing, **73**, 3191-3199 (2010).

[22] W. K. Wong and Z. X. Guo, A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm, International Journal of Production Economics, **128**, 614-624 (2010).

Chi-Jie Lu is an associate professor in the Department of Industrial Management at Chien Hsin University of Science and Technology, Taiwan. He got his Ph.D. in Industrial Engineering and Management from Yuan-Ze University, Taiwan. His research and teaching interests are in the area of data mining, time series forecasting, pattern recognition, and machine vision and inspection. He has published articles in various journals, including Pattern Recognition, Decision Support Systems, International Journal of Production Economics, European Journal of Operational Research, Image and Vision Computing, and Computational Statistics and Data Analysis.

Yuehjen E. Shao is a Professor in the Department of Statistics and Information Science at Fu Jen Catholic University, Taipei, Taiwan. He holds a Ph.D. in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute, New York, U.S.A. His present research interests include statistical process control, quality management, forecasting methods and data mining.

Chin-Chan Li is a graduate student in the Department of Statistics and Information Science at Fu Jen Catholic University, Taipei, Taiwan. His current research interests include machine learning, statistical process control and data mining.